Complete genome sequences of Zika Virus strains isolated from the blood of patients in Thailand (2014) and Philippines (2012).

Ellison,D.W.[1], Ladner,J.T.[3], Buathong,R.[4], Alera,M.T.[1], Wiley,M.R.[3], Hermann,L.[1], Rutvisuttinunt,W[1]., Klungthong,C.[1], Chinnawirotpisan,P.[1], Manasatienkij,W.[1], Melendez,M.C.[2], Maljkovicberry,I.[2], Thaisomboonsuk,B.[1], Ong-ajchaowlerd,P.[1], Kaneechit,W.[1], Velasco,J.M.[1], Tac-An,I.A.[5], Villa,D.[5], Lago,C.B.[1], Roque,V.G. Jr.[6], Akrasewi,P.[4], Plipat,T.[4], Nisalak,A.[1], Srikiatkhachorn,A.[7], Fernandez,S.[8], Yoon,I.K.[9], Haddow,A.D.[3], Palacios,G.F.[3], Jarman,R.G.[2] and Macareo,L.R.[1]

[1] Armed Forces Research Institute of Medical Sciences, Bangkok, Thailand

[2] Walter Reed Army Institute of Research, Silver Spring, MD 20910, USA

[3] Center for Genome Sciences, United States Army Medical Research Institute of Infectious Diseases, 1425 Porter Street, Fort Detrick, Frederick, MD, 21702, USA

[4] Department of Disease Control, Bureau of Epidemiology, Ministry of Public Health, Nonthaburi, Thailand

[5]Cebu City Health Department, Cebu City, Philippines

[6]Department of Health, Manila, Philippines

[7]University of Massachusetts Medical School, Worcester, Massachusetts, USA

[8]United States Army Medical Materiel Development Activity, Frederick, MD, USA

[9]International Vaccine Institute, Seoul, Republic of Korea.

23

ABSTRACT

ZIKV is an arbovirus and member of the family *Flaviviridae* it is transmitted throughout Africa and Asia and is currently causing an outbreak in South America. Here we present the complete genome sequences of two Zika Virus (ZIKV) strains, Zika virus/H.sapiens-tc/THA/2014/SV0127-14 and Zika virus/H.sapiens-tc/PHL/2012/CPC-0740, isolated from the blood of patients collected in Thailand, 2014 and the Philippines, 2012. Sequencing and phylogenetic analysis showed that both strains belong to the Asian lineage.

Key words: Zika, *Flaviviridae*, *Flavivirus*, whole genome sequence

Zika Virus (ZIKV) has garnered worldwide attention as researches have linked the virus to an increase in microcephaly cases during the current outbreak in South America (1). Once thought to cause mild infections, ZIKV is now the subject of intensive worldwide research collaborations and efforts. ZIKV is a single-stranded positive-sense RNA arbovirus and a member of the *Flaviviridae* family, which includes dengue, yellow fever, St. Louis encephalitis, Japanese encephalitis and West Nile viruses (2). The ZIKV genome, approximately 11 kilobases in length, is similar in its arrangement to other members of *Flaviviridae* containing 5' and 3' untranslated regions flanking a single open reading frame (ORF). The 5' and 3' untranslated regions are thought to be important for host interaction, viral replication, and pathogenicity. The ORF codes for three structural proteins: the capsid (C), premembrane/membrane (prM) and envelope (E) and seven nonstructural proteins (NS1, NS2A, NS2B, NS3, NS4A, 2K, NS4B and NS5), which are responsible for viral replication and assembly (3). Previous phylogenetic analysis based on the nucleotide sequences of ZIKV indicated two major lineages: African and Asian (2).

These two ZIKV isolates were obtained by intrathoracic inoculation of ZIKV real-time RT-PCR positive patient's serum samples into *Toxorhynchites splendens* mosquitoes followed by inoculation of mosquito derived C6/36 cells (4, 5). Viral RNA was extracted using the Direct-zol RNA extraction kit (Zymo Research), converted to cDNA using SuperScript III (Invitrogen) and amplified using sequence-independent single-primer amplification (6) combined with primers for rapid amplification of cDNA ends (7). Sequencing libraries were constructed using PrepX ILM 32 DNA Library Kit (Wafergen) and sequenced by using Illumina NextSeq platform (2 x 151 bp). Adaptors and primers were clipped from the sequence reads using Cutadapt version 1.21 (8) and low-quality reads/bases were filtered using Prinseq-lite version 0.20.4 (1) ZIKV

consensus genomes were assembled using Ray Meta (9), Bowtie2 v. 2.0.6 (10) and Samtools v. 0.1.18 (11).

The complete genome sequence with a total length of 10,807 nucleotides (nt) containing the 5' (107 nt) and 3' (428 nt) UTRs and one ORF (10,272 nt) were obtained from both isolates (12). Our sequences and other ZIKV sequences from GenBank were used in maximum-likelihood phylogenetic analysis by using PhyML 3.1 (13) with GTR+G model (–lnL 33630.580); the tree revealed that these two isolates belong to the Asian lineage and are closely related to 2015 Brazilian isolates. This confirms our previous analyses based of NS5 genes of these two isolates (4, 5). Data suggests that Asian lineage isolates are 95-98% identical on the nt level and 98-99% identical on the amino acid level. Nevertheless, some substitutions were found within primer/probe binding sites in the genomes of these two isolates sequences. The two mismatches found on Zika virus/H.sapiens-tc/PHL/2012/CPC-0740 genome for primer 835 (residue 1/23), and for probe 1107 (residue 1/31) binding sites (14). There are two mismatches found on Zika virus/H.sapiens-tc/THA/2014/SV0127-14 genome for primer 911c (residue 21/22), and for probe 1107 (residue 19/31) binding site (14). Further ZIKV genome studies and comparisons will not only elucidate factors involved in the virulence and pathogenicity of ZIKV, but will also lend insight into the evolution of this virus and help with vaccine design.

Nucleotide sequences accession numbers. The assembled complete genome sequences of the Zika virus/H.sapiens-tc/THA/2014/SV0127-14 and Zika virus/H.sapiens-tc/PHL/2012/CPC-0740 isolates were submitted to GenBank under the accession numbers KU681081 and KU681082, respectively. The versions described in this paper are the third versions.

ACKNOWLEDGEMENTS

All military service members or employees of the U.S. Government and this work was prepared as part of their official duties. Title 17 U.S.C. §105 provides that 'Copyright protection under this title is not available for any work of the United States Government.' Title 17 U.S.C. §101 defines a U.S. Government work as a work prepared by a military service member or employee of the U.S. Government as part of that person's official duties. This work is supported by the Global Emerging Infections Surveillance and Response System, a Division of the Armed Forces Health Surveillance Center. Work at USAMRIID was funded by DTRA/JSTO under project CB10246. The views expressed in this article are those of the author and do not necessarily reflect the official policy or position of the Department of the Army, the Department of Defense, nor the U.S. Government.

REFERENCES

1. **Calvet G, Aguiar RS, Melo AS, Sampaio SA, de Filippis I, Fabri A, Araujo ES, de Sequeira PC, de Mendonça MC, de Oliveira L, Tschoeke DA, Schrago CG, Thompson FL, Brasil P, Dos Santos FB, Nogueira RM, Tanuri A, de Filippis AM.** 2016. Detection and sequencing of Zika virus from amniotic fluid of fetuses with microcephaly in Brazil: a case study. Lancet Infect Dis. http://dx.doi.org/10.1016/ S1473-3099(16)00095-5.

2. **Haddow AD, Schuh AJ, Yasuda CY, Kasper MR, Heang V, Huy R, Guzman H, Tesh RB, Weaver SC.** 2012. Genetic characterization of Zika virus strains: geographic expansion of the Asian lineage. PLoS Negl Trop Dis. 6(2):e1477. http://dx.doi.org/10.1371/journal.pntd.0001477

3. **Kuno G and Chang GJ**. 2007. Full-length sequencing and genomic characterization of Bagaza, Kedougou, and Zika viruses. Arch Virol. 152(4):687-96. http://dx.doi.org/10.1007/s00705-006-0903-z

4. **Buathong R, Hermann L, Thaisomboonsuk B, Rutvisuttinunt W, Klungthong C, Chinnawirotpisan P, Manasatienkij W, Nisalak A, Fernandez S, Yoon IK, Akrasewi P, Plipat T.** 2015. Detection of Zika Virus Infection in Thailand, 2012-2014. Am J Trop Med Hyg. http://dx.doi.org/10.4269/ajtmh.15-0022.

5. **Alera MT, Hermann L, Tac-An IA, Klungthong C, Rutvisuttinunt W, Manasatienkij W, Villa D, Thaisomboonsuk B, Velasco JM, Chinnawirotpisan P, Lago CB, Roque VG Jr,**

Macareo LR, Srikiatkhachorn A, Fernandez S, Yoon IK. 2015. Zika virus infection, Philippines, 2012. Emerg Infect Dis. (4):722-4. http://dx.doi.org/10.3201/eid2104.141707.

6. **Djikeng A, Halpin R, Kuzmickas R, Depasse J, Feldblyum J, Sengamalay N, Afonso C, Zhang X, Anderson NG, Ghedin E, Spiro DJ**. 2008. Viral genome sequencing by random priming methods. BMC genomics 9:5. http://dx.doi.org/ 10.1186/1471-2164-9-5.

7. **Leguia M, Loyola S, Rios J, Juarez D, Guevara C, Silva M, Prieto K, Wiley M, Kasper MR, Palacios G.** 2015. Full Genomic Characterization of a Saffold Virus Isolated in Peru. Pathogens 4:816-825. http://dx.doi.org/ 10.3390/pathogens4040816.

8. **Martin M.** 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal 17:10-12. http://dx.doi.org/10.14806/ej.17.1.200.

9. **Boisvert, S., Raymond, F., Godzaridis, E., Laviolette, F., and Corbeil, J.** 2012. Ray Meta: scalable de novo metagenome assembly and profiling. Genome biology *13*, R122. http://dx.doi.org/ 10.1186/gb-2012-13-12-r122.

10. **Langmead B, Salzberg SL**. 2012. Fast gapped-read alignment with Bowtie 2. Nature methods 9:357-359. http://dx.doi.org/ 10.1038/nmeth.1923.

11. **Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R.** 2009. The sequence alignment/map format and SAMtools. Bioinformatics. 25(16):2078-9. http://dx.doi.org/ 10.1093/bioinformatics/btp352.

12. **Ladner JT, Beitzel B, Chain PS, Davenport MG, Donaldson E, Frieman M, Kugelman J, Kuhn JH, O'Rear J, Sabeti PC.** 2014. Standards for Sequencing Viral Genomes in the Era of High-Throughput Sequencing. mBio 5:e01360-01314. http://dx.doi.org/ 10.1128/mBio.01360-

13. **Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O.** 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Systematic biology. 59(3):307-21. http://dx.doi.org/10.1093/sysbio/syq010.

14. **Lanciotti RS, Kosoy OL, Laven JJ, Velez JO, Lambert AJ, Johnson AJ, Stanfield SM, Duffy MR.** 2007. Genetic and serologic properties of Zika virus associated with an epidemic, Yap State, Micronesia. Emerg Infect Dis. (8):1232-9. http://dx.doi.org/10.3201/eid1408.080287.